

AD-A171 888

DIDA - DYNAMIC IMAGE DISPARITY ANALYZER PHASE I(U) AIR  
FORCE WRIGHT AERONAUTICAL LABS WRIGHT-PATTERSON AFB OH  
L A IAMBURINO AUG 85 AFMAL-1R-85-1011

1/1

UNCLASSIFIED

F/G 9/2

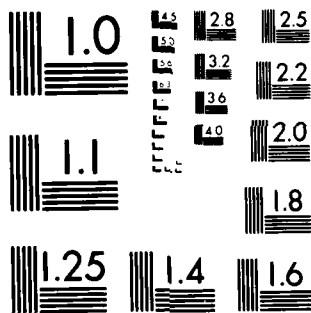
NL

END

DATE

FILED

10-86



XEROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

2

## AD-A171 080

DIDA - DYNAMIC IMAGE DISPARITY ANALYZER

Dr. Louis A. Tamburino

Advanced Systems Research Group  
Information Processing Technology Branch



August 1985

Final Report for Period March 1981 - November 1983

DTIC  
ELECTE  
AUG 13 1986  
S B D

Approved for public release; distribution unlimited.

DTIC FILE COPY

AVIONICS LABORATORY  
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES  
AIR FORCE SYSTEMS COMMAND  
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

86 8 13 015

# NOTICE

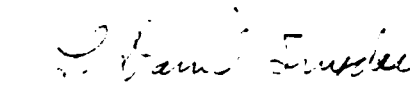
When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Office of Public Affairs (ASD/PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

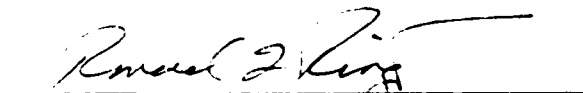


LOUIS A. TAMBURINO  
Project Engineer



L. DANIEL SNYDER, Acting Chief  
Information Processing  
Technology Branch  
System Avionics Division

FOR THE COMMANDER



RONALD L. RINGO, Acting Chief  
System Avionics Division  
Avionics Division

"If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFWAL/AAAT, W-PAFB, OH 45433 to help us maintain a current mailing list".

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFWAL-TR-85-1011		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION AFWAL/AAAT-3	6b. OFFICE SYMBOL (If applicable) AFWAL/AAAT-3	7a. NAME OF MONITORING ORGANIZATION AFWAL/AAAT-3	
6c. ADDRESS (City, State and ZIP Code) Air Force Wright Aeronautical Laboratories (AFSC) Wright-Patterson AFB, OH 45433		7b. ADDRESS (City, State and ZIP Code) Air Force Wright Aeronautical Laboratories (AFSC) Wright-Patterson AFB, OH 45433	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State and ZIP Code)		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO.	PROJECT NO.
		TASK NO.	WORK UNIT NO.
11. TITLE (Include Security Classification) DIDA-DYNAMIC IMAGE DISPARITY ANALYZER		PE62204F	2003
		06	51
12. PERSONAL AUTHOR(S) Louis A. Tamburino			
13a. TYPE OF REPORT Final Report	13b. TIME COVERED FROM Mar 81 TO Nov 83	14. DATE OF REPORT (Yr., Mo., Day) August 1985	15. PAGE COUNT 45
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
X	X	X	
		Image Understanding, Interest Operators, Optical Flow Gradient Statistics, Image Matching.	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>This report summarizes the work accomplished under Phase I of the Dynamic Image Disparity Analyzer (DIDA) program. The DIDA program was initiated to investigate the problems associated with real-time or dynamic disparity analysis. Image disparity analysis is the determination of geometrical differences between two or more images caused by binocular parallax, camera motion, object motion, or some combination of these. The disparity field can provide important 3-D information about structure and motion which is impossible or very difficult to derive from a single image. This report outlines DIDA related efforts; however, its main focus is on in-house achievements. A gradient statistics formalism was developed in order to provide a unified treatment of interest operators<sup>a</sup> for selecting matchable points in different images. This formalism was also used to develop several new local operators for characterizing orthogonal and alignment aspects of local gradients. Computing complete image-to-image disparity maps in real-time is currently an unsolved problem. An adaptive image matching approach is</p>			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> OTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Louis A. Tamburino		22b. TELEPHONE NUMBER (Include Area Code) (513) 255-7648	22c. OFFICE SYMBOL AFWAL/AAAT-3

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

19. (continued)

presented which addressed the complete mapping problem.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

FOREWORD

This report summarizes the work accomplished under Phase I of the Dynamic Image Disparity Analyzer (DIDA) program. This in-house effort was conducted by the Advanced Systems Research Group (AAAT-3), Information Processing Technology Branch (AAAT), Avionics Laboratory (AFWAL/AA), Wright-Patterson AFB, Ohio, under Project 2003, Task 06, Work Unit 51.

The research was performed during the period March 1981 to November 1983.

DTIC  
ELECTE  
AUG 13 1986  
S B



✓	
APPROVED FOR RELEASE	
DATE: 10/1/82	
BY: [Signature]	
A-1	

## TABLE OF CONTENTS

SECTION	PAGE
I INTRODUCTION	1
II TUTORIAL	4
1. Differencing Techniques	4
2. Spatial/Temporal Gradient Analysis	5
3. Matching Techniques	8
4. Interest Operators	10
5. Relaxation Formulation	11
6. Implementation	13
III DIDA RESEARCH PROGRAM	14
IV GRADIENT STATISTICS	19
1. Discrete Gradient Vectors	19
2. Improved Interest Measure Algorithms	21
3. Interpretation of Gradient Histograms	24
4. ABC's of Directional Weighting Functions	26
5. Orientation Aspects	28
6. New Interest Measures	32
7. Section Summary	33
V ADAPTIVE IMAGE MATCHING ALGORITHM	35
REFERENCES	39



## SECTION I

### INTRODUCTION

This report summarizes the work accomplished under Phase I of the Dynamic Image Disparity Analyzer (DIDA) program from March 1981 to November 1983. Image disparity analysis is the determination of geometrical differences between two or more images caused by binocular parallax, camera motion, object motion, or some combination of these. Disparities between images can be represented as a vector field mapping one image into the corresponding points of the other image. The disparity field can provide important 3-D information about structure and motion which is impossible or very difficult to derive from a single image.

The goal of the DIDA program was to develop advanced real-time systems, as depicted in Figure 1. Dynamic image analysis embraces two distinct areas: (1) the correspondence problem and (2) interpretation. This report (and most research in the literature) deals with the first area, the correspondence problem of associating corresponding structures in different images; i.e., the determination of the disparity field. The interpretation problem is concerned with transforming measured disparity fields derived from 2-D images into physical parameters related to 3-D structure and motion of objects. Interpretation procedures are poorly understood.

Current papers on the interpretation problem assume solutions to the correspondence problem which unfortunately are not always possible. Theory and formalism are presently in a state of flux. Journal articles are usually less than ten years old in this area, and there are no comprehensive review articles or text books. One encounters major difficulties approaching an interpretation problem with no a priori knowledge. The basic reason for the difficulties is that the general problem deals with non-linear equations of enormous complexity. These equations have multiple answers and require numerical analysis to achieve approximate solutions. These numerical methods may converge to a wrong answer and are sensitive to initializing parameters. Additional constraints and other a priori knowledge are often utilized to simplify both disparity analysis and the difficult non-linear interpretation problems. It is too early now to anticipate the final formulations of and practical solutions to interpretation problems.

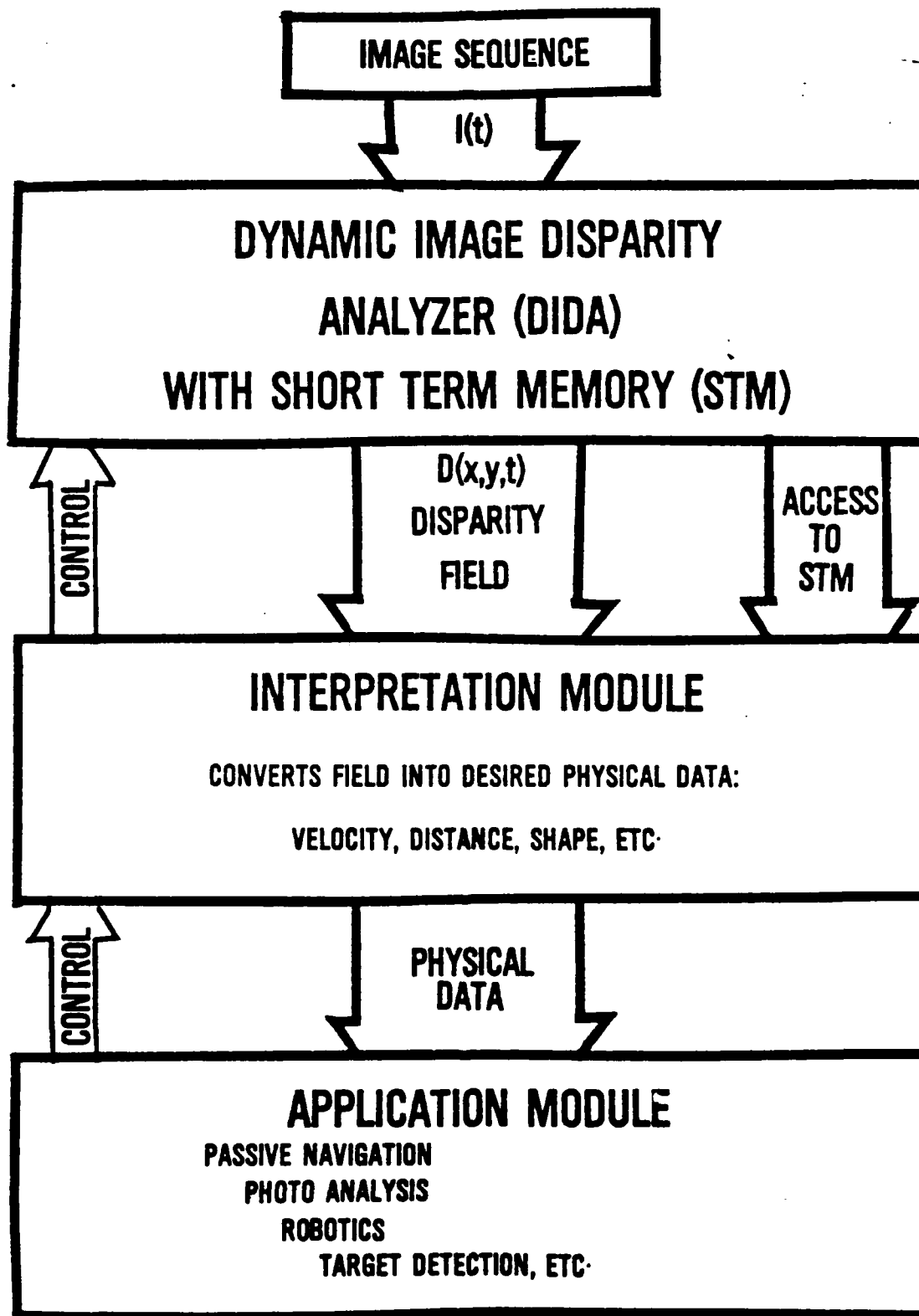


Figure 1. Dynamic Image Analyzer

The type of information derivable from a disparity field includes depth, surface orientation, object boundaries, camera motion, and object motion. Derivation of this physical data from sensor inputs is fundamental to many military applications, including automatic target detection, recognition, and tracking, and passive navigation with stereo-motion imagery. Currently, there is much research on the optical flow (continuous field of disparities) problem which seeks to derive an approximate model of the earth's surface from a sequence of images taken from a flying sensor. Even a crude surface model provides passive ranging which is extremely useful in determining whether a blob in the image is the proper size for selected targets. At the 1983 DARPA Image Understanding Workshop nearly half the papers dealt with the optical flow problem, which is one of the relatively easy interpretation problems.

The various successful approaches to the correspondence problems are reviewed and discussed in Section II. One conclusion of this review is that these approaches do not form an adequate basis for developing a general purpose real-time high density disparity analyzer. In the initial phase of the DIDA program, research efforts were mainly concerned with developing improved algorithms for disparity analysis. The overall DIDA program plan and other research done in support of it are discussed in Section III. Theoretical results derived by the author are presented in Sections IV and V.

## SECTION II

## TUTORIAL

The ideal solution to the correspondence problem is found in biological vision systems which are still a mystery. Nature uses a highly parallel network of simple processors which communicate with their nearest neighbors. The system extracts low level features from receptor outputs in the retina and weaves these local features into a global perceptive tapestry. The determination of a universal disparity algorithm is a very difficult and unsolved problem. These approaches have been developed which can extract disparity information within their respective domains. These approaches are referred to as (1) differencing techniques, (2) spatial-temporal gradient analysis, and (3) feature matching.

## 1. DIFFERENCING TECHNIQUES

Differencing techniques provide for a point by point determination of significant changes in a sequence of images (Reference 1, 2, 3, and 4). Differencing requires registered images and is therefore not suitable for moving camera applications. Early techniques were based on thresholding simple temporal differences:

$$I(X,Y,T) - I(X,Y,T - \Delta T) > \epsilon \quad (1)$$

Current techniques are more elaborate neighborhood functions such as

$$\frac{[(S_C + S_R)^2 + (M_C - M_R)^2]^2}{S_C * S_R} > \epsilon \quad (2)$$

where M and S = mean and variance for sample areas from current (C) and reference (R) image frames. Thresholding results in binary difference images which essentially act as motion detectors. Disparity fields are not measured directly. In addition to image registration, occlusion problems and boundary definition require sophisticated algorithms to track, integrate, and interpret a sequence of binary images. Application of this approach is suited to initial phases of dynamic image analysis for initial estimates of velocity field clusters and segmentation estimates.

## 2. SPATIAL/TEMPORAL GRADIENT ANALYSIS

The second approach combines the intensity gradient and temporal gradient at each pixel in the image; i.e.,  $\vec{\nabla}I$  and  $\delta I/\delta t$  (References 5, 6, and 7). The basic assumption made in this approach is that the brightness of a displaced patch of a moving object remains constant:

$$I(X,Y,t) = I(X+\Delta X, Y+\Delta Y, t+\Delta t) \quad (3)$$

where the patch undergoes a displacement of  $X, Y$  in  $t$ . With this assumption, one can derive a velocity constraint which relates the velocity to changes in intensity over both time and space:

$$\frac{\delta I}{\delta X} \frac{\delta X}{\delta t} + \frac{\delta I}{\delta Y} \frac{\delta Y}{\delta t} + \frac{\delta I}{\delta t} = 0 \quad (4)$$

where higher order derivatives in the Taylor expansion are suppressed in the limit  $\Delta t \rightarrow 0$ . The gradient and velocity vectors are given by

$$\vec{\nabla}I = [\delta I/\delta X, \delta I/\delta Y] \quad (5)$$

and

$$\vec{V} = [\delta X/\delta t, \delta Y/\delta t] \quad (6)$$

where

$$\Delta X/\Delta t \rightarrow \delta X/\delta t, \quad \Delta Y/\Delta t \rightarrow \delta Y/\delta t, \quad \text{and} \quad \Delta I/\Delta t \rightarrow \delta I/\delta t.$$

An analysis of higher order terms in the Taylor expansion was made in a study of error reduction techniques by Thompson and Kearney (Reference 7).

The basic constraint equation can be rewritten as a vector equation:

$$\vec{\nabla}I \cdot \vec{V} + \delta I/\delta t = 0. \quad (7)$$

Note that the constraint equation gives only the component of the velocity vector (or optical flow) which is parallel to the gradient vector. The optical flow field, therefore, cannot be computed at a point independent of neighboring points. Additional constraints are necessary to find the perpendicular component of the velocity vector. Two reported ways of augmenting the basic constraint equation are the clustering approach and the using of a smoothness constraint.

The clustering approach is similar to a Hough transform and suitable for segmenting a sequence of images with multiple rigid body translations and in which there is no camera motion. The unknown velocity component  $v_{\perp}$  lies on a constant line perpendicular to the  $v_{||}$  component. Several constraint lines are shown in Figure 2. In general, the constraint lines of neighboring points will not be colinear and therefore will intersect. Over the cluster of intersecting points some average is made to estimate  $v_{\perp}$ . A pre-smoothing operation to remove noise and high frequency texture variations generally improves velocity estimates. Thompson and Barnard (Reference 8) described a "pseudo inverse" method which is equivalent to least square error analysis and similar to clustering analysis. These methods have been successfully tried on real images. They involve relatively complex computations and cannot deal with rotational motion which generates a continuum of velocity directions.

A smoothness constraint is obtained from a measure of the departure from smoothness in the velocity flow, such as:

$$E_c = [(\delta U / \delta X)^2 + (\delta U / \delta Y)^2] + [(\delta V / \delta X)^2 + (\delta V / \delta Y)^2] \quad (8)$$

where  $u, v$  are the  $x, y$  velocity components (Reference 5 and 6). Let

$$E_b = \frac{\delta I}{\delta X} U + \frac{\delta I}{\delta Y} V + \frac{\delta I}{\delta t} \quad (\text{basic velocity constraint}) \quad (9)$$

Constraint equation are obtained by minimizing

$$E = \int \int (\alpha E_c^2 + E_b^2) dX dY \quad (10)$$

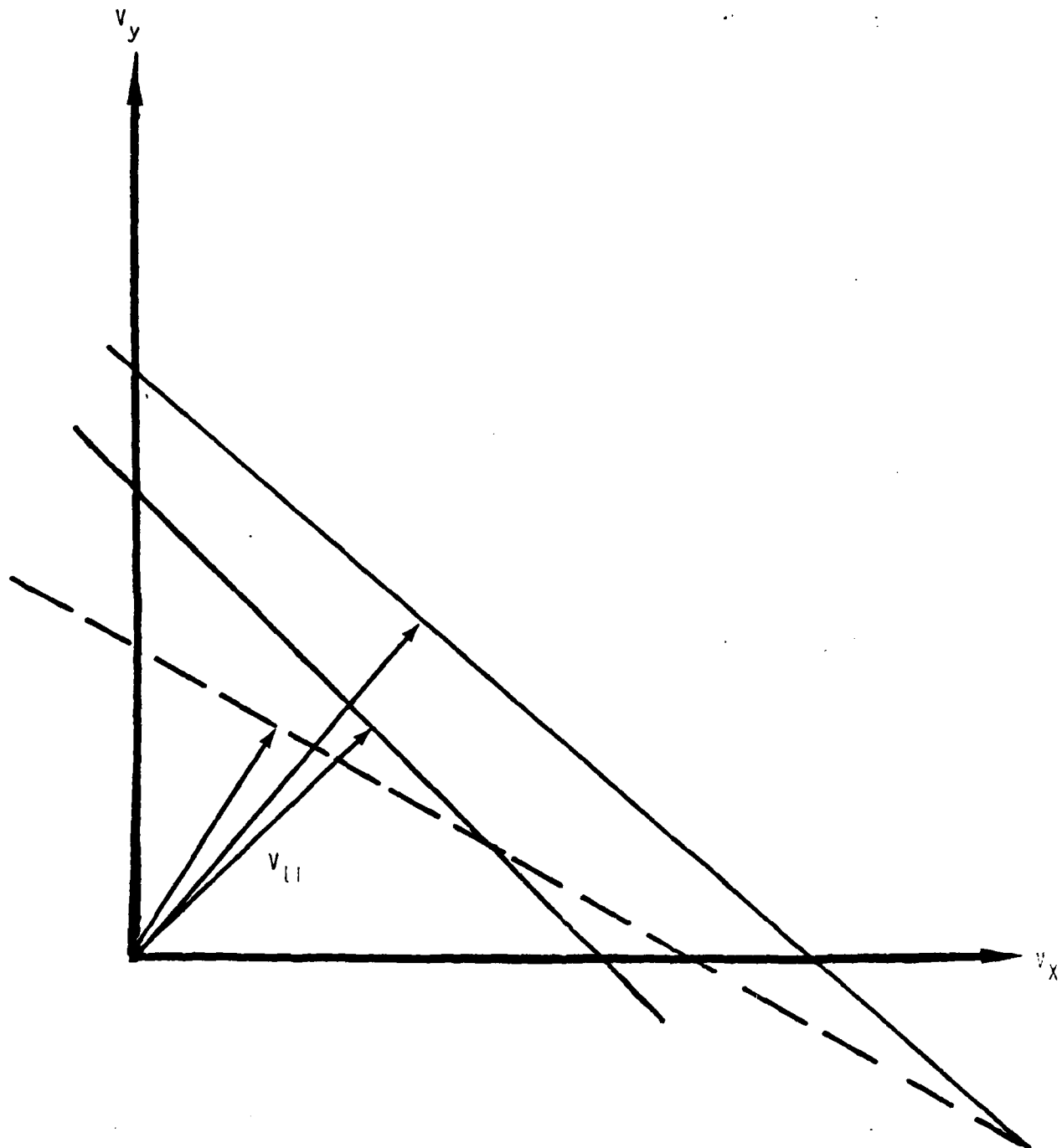


Figure 2. Constraint Lines

using the calculus of variation. This particular measure results in a set of iterative equations:

$$U^{n+1} = \bar{U}^n - (\delta I / \delta X) \cdot \bar{E}_b / (\alpha^2 + G^2) , \quad (11)$$

$$V^{n+1} = \bar{V}^n - (\delta I / \delta Y) \cdot \bar{E}_b / (\alpha^2 + G^2) , \quad (12)$$

$$G^2 = (\delta I / \delta X)^2 + (\delta I / \delta Y)^2 , \quad (13)$$

where the neighborhood averages indicated by bars result in a "fill-in" feature for uniform regions. That is to say, regions with zero gradients have no local information to constrain velocity. In this case, velocities are taken to be neighborhood averages. Eventually, velocities in non-uniform regions or border regions will propagate inwards and fill in ambiguous regions. This corresponds to obtaining a solution to the Laplace equation with given boundary conditions. So far, this approach has been applied mainly to synthetic images. It works best when the gradient is not too small and varies in direction from point to point. The smoothness assumption breaks down at discontinuities which may prove to be a major flaw in this approach. Heuristics have been proposed for establishing motion boundaries using maximum velocity and orientation similarity.

The foregoing approach was reviewed for two reasons: first, to present a technique which seeks to find a disparity vector for each pixel and secondly, to dispel any inference that the computation of the optical field is a closed problem.

### 3. MATCHING TECHNIQUES

There are many different approaches to image matching, depending on what is to be matched and how it is to be matched. The raw data of image analysis is a matrix of image intensities. Hence, intensity matching using cross-correlation has been extensively studied and utilized. In image analysis, raw intensity data is transformed into various representations characteristic of one or more features. Features are derived from the spatial variations and relationships of the distribution of features or attributes such as position, size, orientation, color, adjacencies, etc. Symbolic matching involves less matching ambiguity because complex



features are unique. Another advantage is that they are more invariant to substantial and arbitrary global changes and orientation. A typical strategy of many symbolic matching algorithms is the coarse-fine strategy in which one selects and pairs the most obvious matching features initially. Next, these initial matches are used to find approximate coordinate transformations and scaling factors. Using this transformation, the remaining segments are modified and matched, the transformation then refined, etc. Thus, one can see that symbolic matching can work with large rotations which foil other approaches.

Simple or primitive features such as points, micro-edges, blobs, and texture measures are local functions of the intensity values over a small neighborhood of pixels (picture elements). Computation of local features is amenable to pipeline and parallel processing techniques. An image matching algorithm is outlined in Section V which features an adaptive blob-matching approach for deriving dense disparity fields. Simple features, such as points, have multiple possible matches and require additional measures to reduce ambiguity. With  $n$  elements selected in each image, there are  $n!$  possible one-to-one mappings. These measures are illustrated by the label relaxation technique for matching point sets, which is described below. Point matching will be discussed in some detail because of its relationship to work done on "interest measures" under the DIDA in-house research effort.

The following point matching algorithm, which assumes limits for the amount of displacement and rotation between images, has been described by Thompson and Barnard (Reference 8 and 9):

- a. Use "Interest Operator" to locate unique set of points in each image.
- b. Assign labels and probabilities to candidate matches.
- c. Use label relaxation technique to increase local consistency and reduce matching ambiguity.

## 4. INTEREST OPERATORS

There is no generally preferred local method for identifying the same set of points in multiple images. The matchability of any point depends on the intensity values of neighboring points. A uniqueness measure for each pixel is a function of intensity values over a small window about each pixel. Points with the highest measure in their local region are selected as "interesting" points which have the highest likelihood of being matched. One early measure is simply the statistical variance of image intensities within the window:

$$\text{VAR} = \text{MEAN} (I(i,j) - \bar{I}(i,j))^2 \quad (14)$$

where

$$\bar{I} = \text{MEAN} (I(i,j)).$$

In order to improve localization of selected points, other operators were developed which are more sensitive to the crossings of multiple linear features. Two popular window measures are based on the means of the square of intensity differences between adjacent pixels in four different directions:

$$\begin{aligned} D(1) &= \text{MEAN} (I(i,j) - I(i+1,j))^2 \\ D(2) &= \text{MEAN} (I(i,j) - I(i,j+1))^2 \\ D(3) &= \text{MEAN} (I(i,j) - I(i+1,j+1))^2 \\ D(4) &= \text{MEAN} (I(i+1,j) - I(i,j+1))^2 \end{aligned} \quad (15)$$

The Moravec interest measure (Reference 10 and 11) is the directed variance:

$$\text{DIRVAR} = \text{MIN} [D(1), D(2), D(3), D(4)] \quad (16)$$

The second measure is the Hannah edge variance (References 12 and 13):

$$\text{EVAR} = \text{VAR} * \text{MIN} \left[ \frac{D(1)}{D(2)}, \frac{D(2)}{D(1)}, \frac{D(3)}{D(4)}, \frac{D(4)}{D(3)} \right] \quad (17)$$

Both the Moravec and Hannah interest measures seek out the direction associated with minimum difference. If the minimum is large, the differences in all directions are large; hence, the features selected tend to be associated with multi-directional attributes. This in turn insures that the feature is localized and not dominated by a difference due to a single linear feature.

Point selection or interest operators can be defined as local maxima of an interest measure. A generalized treatment of point selection operators will be presented below based on gradient statistics. It will be shown that the computation of the Moravec and Hannah operators can be improved and simplified. A new set of point selection operators are presented for future experimentation and evaluation.

## 5. RELAXATION FORMULATION

The relaxation labeling process can be broken down into two phases: initialization and iteration. The discussion follows that of W. Thompson (Reference 8). The initialization phase begins after a set of unique points have been selected from each image. The problem is to match the points in one image called the reference image with the corresponding points in the second image. For each point "i" in the reference image, one constructs a label list  $L^i = [\vec{l}_{ij}, l_i^*]$ , where  $\vec{l}_{ij} = \vec{x}_i - \vec{x}_j$  is the vector displacement between point "i" in the reference image and point "j" in the second image. Each  $\vec{l}_{ij}$  is a candidate match or disparity vector between two points.  $l_i^*$  is a special no-match label for point "i".

The next step is to assign initial probabilities  $p(l_{ij})$ ,  $p(l_i^*)$  to each label based on local similarity measures (abbreviated  $p_{ij}$ ,  $p_i^*$ ). The initial probabilities  $p^0(l_{ij})$  are based on correlations between a small reference window from one image centered on  $\vec{x}_i$  and a window from the second image centered at  $(\vec{x}_i + \vec{l}_{ij})$ . The sum of the squares of the differences between windows has been used as the measure for the amount of correlation. Many potential matches can be eliminated if an upper limit on the magnitude for all the disparity vectors is assumed.

A consistency measure,  $q_{ij}^t$ , is the sum of probabilities  $p_{kv}^t$  for all labels  $l_{kv}$  closely aligned with  $l_{ij}$  and in the same neighborhood:

$$q_{ij}^t = \sum \sum p_{kv}^t \cdot e_{kv} \quad (18)$$

where  $e_{kv}$  is equal to one if both of the following conditions are satisfied and equal to zero otherwise.

(1) Condition for point "i" to be near to point "k":

$$\max [|x_i - x_k|, |y_i - y_k|] \leq R \text{ (constant)}. \quad (19)$$

(2) Condition for  $\vec{l}_{ij}$  to be closely aligned to  $\vec{l}_{kv}$ :

$$|\vec{l}_{ij} - \vec{l}_{kv}| < \epsilon \text{ (constant)}. \quad (20)$$

The consistency measures are used in turn to update the label probabilities.

Let

$$\hat{p}_{ij}^{t+1} = p_{ij}^t * (A + B * q_{ij}^t) \quad (21)$$

and

$$\hat{p}_i^{t+1*} = p_i^{t*} \quad (22)$$

represent the new unnormalized probabilities where A and B are adjustable gain parameters (typically A = 0.3 and B = 3). To normalize, let

$$p^{t+1} = \hat{p}^{t+1} / \left( \sum_{\text{ALL}} \hat{p}^{t+1} \right). \quad (23)$$

Note that the probability of the no match label is affected only by this normalization step.

## 6. IMPLEMENTATION

Current implementation of matching systems focuses mainly on stereo systems, either binocular or motion (monocular) stereo. All these approaches are characterized by point or edge matching between images taken at approximately the same time and perspective. In stereo matching, the possible disparities are tightly constrained. Here the matching problem requires a dense grid of points making the matching problem considerably more difficult than determination of distance. Some systems concerned with automatic or robotic navigation will trade off density for accuracy. Sub-pixel accuracy is sought using point matching and cross correlation. At the time of this report, there are no complete real-time systems. However, a near real-time prototype matching system is being developed at MIT, which produces a 16 by 16 array of depth measurements every 15 seconds from a vidicon camera. For more information, see the excellent survey article by Barnard and Fischer, "Computational Stereo From an IU Perspective" in the Proceedings: DARPA Image Understanding Workshop (April 1981).

Optical techniques are not included in this survey.

### SECTION III

#### DIDA RESEARCH PROGRAM

Dynamic and 3-D image analysis is a rapidly emerging research area. Real-time systems, however, are slowly emerging because current algorithms have limitations. This area of research is difficult because it includes automatic image understanding, non-linear adaptive scene analysis, enormous data manipulation problems, and knowledge base design problems beyond current methodology. This research, which is basic to developing robust computer vision, will certainly draw upon and enhance our understanding of biological vision systems.

The in-house DIDA program was initiated at AFWAL/AAAT in the early part of 1981. DIDA is an acronym for Dynamic Image Disparity Analyzer. The long term objectives of the DIDA program were to focus research in the area of real-time vision systems and to develop requisite in-house expertise and processing facilities. The initial objective was to conduct research applicable to the development of a real-time disparity analyzer. Figure 3 depicts the DIDA Program Plan. This technical report summarizes the work accomplished in Phase I of the program. The DIDA program, if support , would stimulate important main stream research and contribute to advanced image understanding technology for: robotics, passive navigation, photo analysis, target detection, recognition, and tracking, and multi-sensor management techniques.

The work in Phase I was accomplished by AFWAL/AAAT personnel, two AFIT students, and an LDF contracted research effort at the University of Minnesota. The DIDA team roster is shown in Figure 4. Mr. C. Wagner assisted in planning and enhancing the AAAT DIDA laboratory facilities which was to be merged eventually in 1983 with the current laser vision laboratory at AAAT. A follow-on research effort at the U. of Minnesota was funded by the Air Force Office of Scientific Research in 1983. The remainder of this section will briefly summarize the documentation generated in Phase I. Sections IV and V present research results by the author which are not included in the other DIDA reports mentioned below.

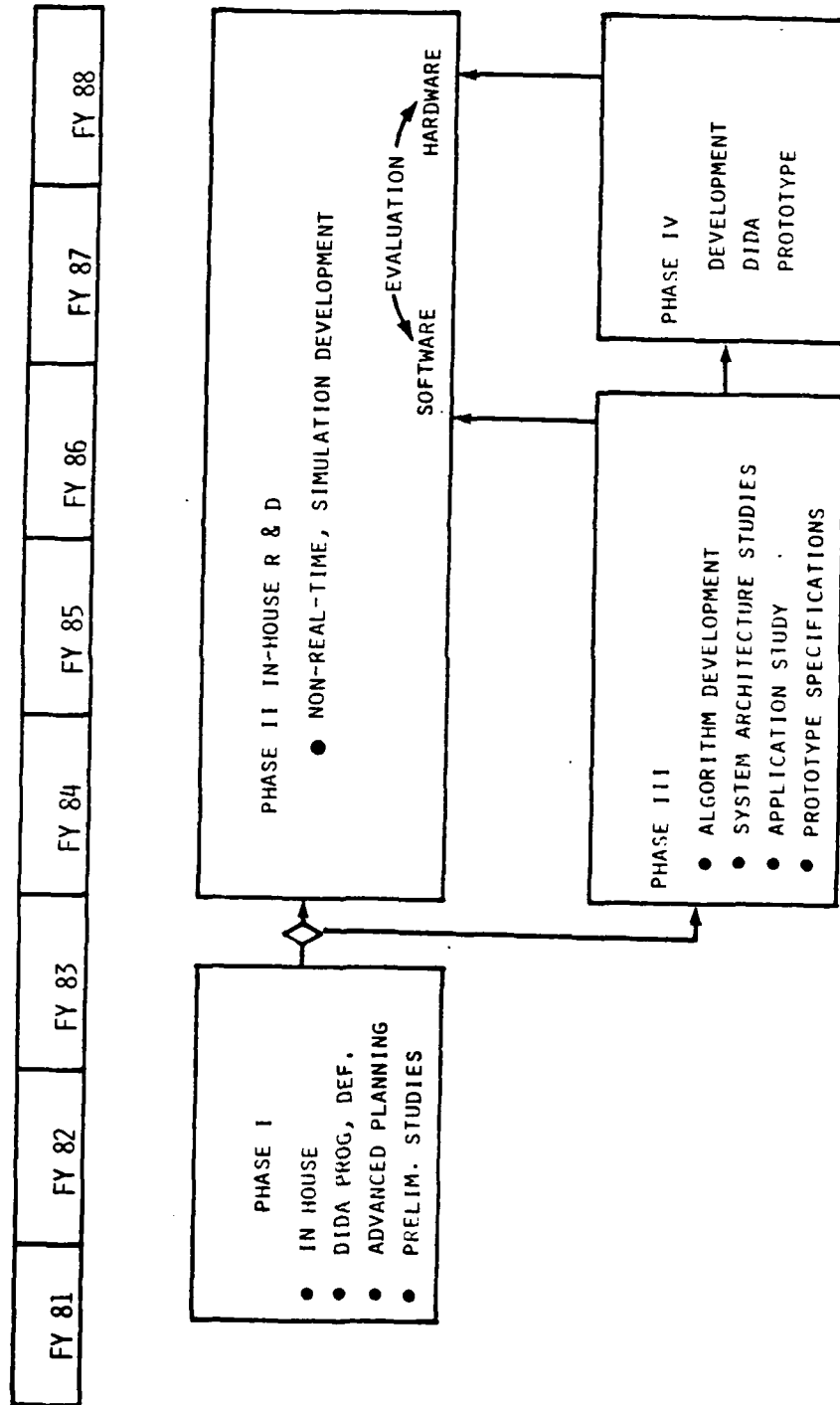


Figure 3. DIDA Program Plan

AFWAL/AAAT	Dr. Louis A. Tamburino Mr. Charles E. Wagner	(Project Engineer) (Part Time 1981)
AFIT	Capt. Franklin D. Cooper Major Michael S. Gaydeski	(1981) (1982)
U. Minnesota	Dr. William B. Thompson Mr. Joseph K. Kearney	(1981-1982) (1981-1982)

Figure 4. DIDA Roster



Capt. F. Cooper's Master's thesis entitled "Disparity Analysis of Time-Varying Imagery" (AFIT/GE/EE/810-1) attempted to develop a software simulation package to evaluate point selection and point matching algorithms. In the six month period allotted to AFIT students, Capt. Cooper managed to produce experimental photographs of his computer simulation experiments and obtain preliminary results. While these preliminary experiments provide instructive insights into the subtleties of point matching algorithms, more debugging of the software is required before its general acceptance. This work is documented in a 370 page thesis, including software listings. In the course of this work, AAAT acquired for Capt. Cooper's use image data bases from three prominent research centers which will also prove useful in future research.

Major M. Gaydeski's Master's thesis, entitled "Disparity Analysis - Real-Time Determination of Object Motion" (AFIT/FE/EE/82-1) is a survey of image transformations and features which could prove useful in mapping grey level images into binary images via a thresholding function. This survey includes sections on global threshold selection for segmentation, local segmentation, and global local edge coincidence. This literature survey was done in support of the adaptive image matching system discussed in Section V.

The research accomplished at the University of Minnesota by Thompson and Kearney is documented in AFWAL-TR-83-1035 entitled "DIDA - Dynamic Image Disparity Analysis." The objective of this study was to perform the initial analysis required to implement a system for determining disparity values in real-time. The final report describes the results of a study of the feasibility of developing a device for the real-time estimation of motion induced disparity in image sequences. The report describes the nature of the estimation problem and suggests criteria by which methods for estimating disparity can be evaluated. It includes a theoretical analysis of one class of estimation methods and shows how such an analysis can lead to improved performance. The results obtained from a variety of estimation algorithms are demonstrated on a limited sample of dynamic imagery. Finally, preliminary hardware and software requirements are provided for real-time disparity analysis devices. Work under this contract has successfully produced analytical analyses of the intrinsic limitations of several state-of-the-art algorithms. This analysis is being used to modify these algorithms to improve performance. Several

AFWAL-TR-85-1011

hybrid approaches which can combine the strengths of several different existing approaches have been identified.

## SECTION IV

### GRADIENT STATISTICS

#### 1. DISCRETE GRADIENT VECTORS

The original motivation for studying gradient statistics was to develop rotationally invariant and more efficient point selector operators. The gradient statistics formalism which evolved from this study facilitates not only a unified and comprehensive treatment of interest measures, but also the development of other fundamental feature detectors for image understanding systems. Initially, the study developed simplified algorithms for the Moravec and Hannah interest operators; however, it also used the formalism to develop new interest measures which are rotationally invariant. The presentation of results which follows is organized in a manner which reflects the chronological order of this research.

By way of review, a gradient is a vector denoting the magnitude and direction of the maximum rate of change of a scalar. The gradient vector components of a two-dimensional function  $I(x,y)$  are simply the partial derivatives:

$$\text{Grad. } I = (\partial I / \partial x, \partial I / \partial y) \quad (24)$$

or, rewritten using an abbreviated notation,

$$\vec{\nabla} I = (I_x, I_y).$$

If a coordinate rotation is given by

$$x' = x \cos\theta + y \sin\theta \quad (25a)$$

$$y' = -x \sin\theta + y \cos\theta$$

then the gradient vector transformation is

$$I_{x'} = I_x \cos\theta + I_y \sin\theta \quad (25b)$$

$$I_{y'} = -I_x \sin\theta + I_y \cos\theta$$

If, for example,  $\theta = 45^\circ$  then

$$\begin{aligned} x' = \alpha &= (x + y) / \sqrt{2}; \quad I_\alpha = (I_x + I_y) / \sqrt{2} \\ y' = \beta &= (-x + y) / \sqrt{2}; \quad I_\beta = (-I_x + I_y) / \sqrt{2} \end{aligned} \quad (26a)$$

( $\alpha, \beta$  used below to signify  $45^\circ$  rotations).

For this special case, the new coordinates and components are simply the sum and difference divided by  $\sqrt{2}$ .

This brief review of gradients uses the conventional formalism for dealing with analytical functions. How to apply this gradient formalism to digitized images, which are discrete arrays of intensity values, is not obvious. Consider the 2 X 2 array of pixels depicted in Figure 5. Discrete estimates of the partial derivatives of  $I(i,j)$  are given by pixel differences:

$$\bar{I}_x = I(b) - I(a) \text{ or } = I(d) - I(c) \quad (27)$$

$$\bar{I}_y = I(c) - I(a) \text{ or } = I(d) - I(b)$$

Another estimate with respect to the diagonal coordinate system yields

$$I_\alpha = (I(d) - I(a)) / \sqrt{2}$$

$$I_\beta = (I(c) - I(b)) / \sqrt{2} \quad (26b)$$

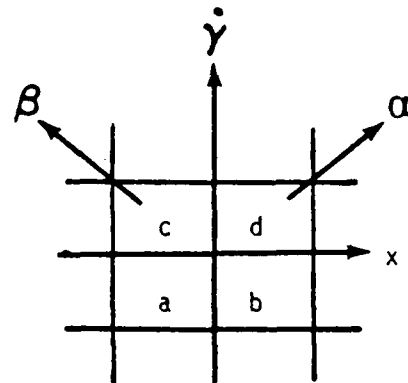


Figure 5. 2 X 2 Array of Pixels

where the factor  $\sqrt{2}$  is the diagonal distance between the diagonal pixels (assuming unit distance between vertical and horizontal neighbors). If one uses Equation 25b to rotate the components  $45^\circ$ , one obtains

$$I_x = (I_\alpha - I_\beta) / \sqrt{2} = [I(b) + I(d) - I(a) - I(c)] / 2 \quad (28)$$

$$I_y = (I_\alpha + I_\beta) / \sqrt{2} = [I(c) + I(d) - I(a) - I(b)] / 2$$

Notice that this discrete approximation for the two partial derivatives differs from Equation 27. These partial derivatives depend on four pixels instead of two. The  $I_x$  partial derivative is the average of two estimates of  $\bar{I}_x$  type, i.e.,  $D - C$  and  $B - A$ ;  $I_y$  is the average of  $C - A$  and  $D - B$ . Although it may appear that matters are slightly more complicated by using Equation 28 in place of Equation 27, it will be shown that subsequent formalisms are not only rotationally consistent, but easier to compute. Definitions 26b and 28 are consistent with the requirement that vector magnitudes be rotationally invariant, i.e.,

$$\rho^2 = I_x^2 + I_y^2 = I_a^2 + I_b^2 \quad (29)$$

The Sobel edge detector, which is one of the most popular edge detectors, is defined by two fundamental measures:

$$S_x = [(c + 2f + i) - (a + 2d + g)]/8$$

$$S_y = [(g + 2h + i) - (a + 2b + c)]/8$$

	g	h	i
	d	e	f
	a	b	c

where a,b,... are used in place of  $I(a)$ ,  $I(b)$ ,....

One can easily see that  $S_x$  equals the average of the four horizontal gradient components (Equation 28) derivable from the four  $2 \times 2$  areas, which fit inside the  $3 \times 3$  area. Similarly,  $S_y$  is also the average of four estimates of the vertical derivatives. Hence, one can think of the Sobel operator as an average over the more fundamental  $2 \times 2$  operators. This local averaging tends to eliminate noise and smooth out edges.

## 2. IMPROVED INTEREST MEASURE ALGORITHMS

If the DIRVAR and EVAR interest measures defined in Section II are reviewed in light of the rotationally consistent gradient definitions defined in this section, several improvements are readily evident. The horizontal and vertical difference statistics (recall Equation 15) can be interpreted as

$$\begin{aligned} D(1) &= \sum [I(i+1,j) - I(i,j)]^2 = \sum \bar{I}_x^2 \\ D(2) &= \sum [I(i,j+1) - I(i,j)]^2 = \sum \bar{I}_y^2 \end{aligned} \quad (30a)$$

where  $\tilde{I}_x$ ,  $\tilde{I}_y$  are the partial derivative estimates are defined by Equation 27. In a similar manner, the diagonal difference statistics can be written as

$$D(3) = \sum [I(i+1, j+1) - I(i, j)]^2 = \sum (\sqrt{2} I_\alpha)^2 \quad (30b)$$

$$D(4) = \sum [I(i+1, j) - I(i, j+1)]^2 = \sum (\sqrt{2} I_\beta)^2$$

where  $I_\alpha$  and  $I_\beta$  are defined in Equation 26b. If the interest operators are to be rotationally invariant, the coordinate system used should make no difference; i.e., the use of  $(I_\alpha, I_\beta)$  and  $(I_x, I_y)$  should be interchangeable. This would mean the diagonal differences  $D(3)$  and  $D(4)$  should be scaled by the factor 1/2 relative to  $D(1)$  and  $D(2)$  in order to account for the  $\sqrt{2}$  distance between diagonal pixels. In other words, by properly scaling differences, one compares the slope of the intensity in four principal directions. This is more meaningful and consistent than comparing just the pixel intensity differences.

Interest measures based on the directional statistics  $D(i)$  would be enhanced and simplified if the following definitions are adopted:

$$D(1) = \sum I_x^2, D(2) = \sum I_y^2 \quad (31)$$

$$D(3) = \sum I_\alpha^2, D(4) = \sum I_\beta^2$$

where  $(I_\alpha, I_\beta)$  are defined by Equation 26b and  $(I_x, I_y)$  by Equation 28. It now follows from invariance of vector magnitudes that the newly defined  $D(i)$  satisfy

$$D(1) + D(2) = D(3) + D(4) \quad (32)$$

hence only three of the  $D$ 's are independent. Equation 32 can be used to solve for the fourth  $D$  which means computations are effectively reduced by 25%.

The three independent statistics which characterize the four  $D$ 's and link them with the formalism to be developed and used in the remainder of this paper are:

$$\begin{aligned} A &= \sum I_\alpha^2 + \sum I_\beta^2 = D(1) + D(2) = D(3) + D(4) \\ B &= \sum I_\alpha^2 - \sum I_\beta^2 = D(3) - D(4) \end{aligned} \quad (33)$$

$$C = \sum I_x^2 - \sum I_y^2 = 2 \sum I_x^2 - A$$

where

$$2 D(i) = [A + C, A - C, A + B, A - B] \quad (34)$$

Substitution of the  $D(i)$  from (Equation 31) into the original definitions for directed variance and edge variance (see Equations 16 and 17) yields

$$DIRVAR = \min [D(i)] = (A - MX)/2 \quad (35)$$

$$EVAR = VAR * \min [D(i)/D(i)] = VAR * (A-MX)/(A+MX) \quad (36)$$

$$\text{where } MX = \max [|B|, |C|]. \quad (37)$$

The selection of the minimum value is now accomplished with one compare operation where originally it took three. Also note that this definition of EVAR requires only one division.

Both the  $(x,y)$  and  $(\alpha,\beta)$  components of the gradient vector are derivable from the same  $2 \times 2$  array of pixels. Since the  $2 \times 2$  arrays fit evenly into rectangular windows, an equal number of components in the four principal directions are obtained within such windows. In the original definitions of the  $D(i)$ , this was not the case (see Equations 15). The original  $D(i)$  were defined as averages to compensate for the variation in the number of diagonal and horizontal or vertical pixel differences. Hence the process of sorting and averaging would require division operations. The new directional statistics (Equation 31) always involve the same number of terms, hence they may be sorted and compared without division by the number of terms.

The improvements in the computation of interest measures demonstrate the advantage to be derived from consistent use of gradient statistics. A more generalized treatment of applications for gradient statistics follows, which will include new interest operators which are rotationally invariant and analytic.

### 3. INTERPRETATION OF GRADIENT HISTOGRAMS

Given a localized set of gradient vectors, how can one extract various interesting image features from this data? Originally, we used gradient histograms as the basis of statistical interpretation. Computation of these gradient histograms requires the direction and magnitude of each vector:

$$\rho = \sqrt{I_\alpha^2 + I_\beta^2} = \sqrt{I_x^2 + I_y^2} \quad (38)$$

$$\theta = \tan^{-1} I_y/I_x = (\tan^{-1} I_\beta/I_\alpha) - \pi/4$$

The angular dimension is quantized into M bins, each spanning  $2\pi/M$  radians. The magnitudes of the vectors falling into each bin are summed

$$h_j = \sum_{i \in \text{Bin}(j)} \rho_i \quad (39)$$

and a normalized set of weights obtained by

$$p_j = h_j / \left( \sum_{k=1}^M h_k \right) \quad (40)$$

Figure 6 depicts some idealized histograms and corresponding examples of localized window images. Histograms may be characterized by various parameters which ideally correspond to meaningful features in the localized images. One example of a useful parameter is entropy,

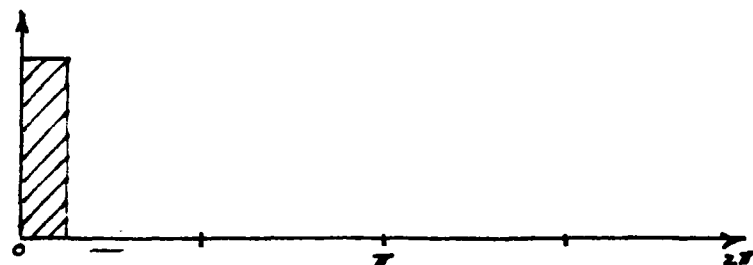
$$E = - \sum_{k=1}^M p_k \ln(p_k) \quad (41)$$

which equals zero for a single edge and equals the maximum value ( $\ln M$ ) for uniform histograms. The entropy measure does not distinguish between histograms b and c which have equal entropy.

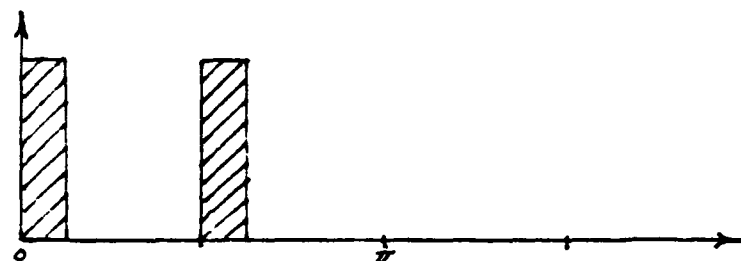
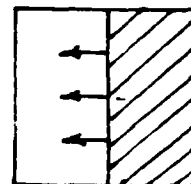
In interest measures, the orthogonal aspects of b, d, and e are significant. The need to distinguish histograms with large orthogonal aspects leads to the following concept of directional weighted averages for each bin:

$$w_I = \sum_{k=1}^M p_k |\sin(\theta_k - \theta_I)| \quad (42)$$

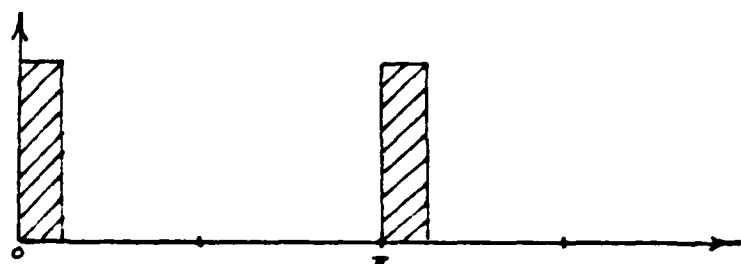
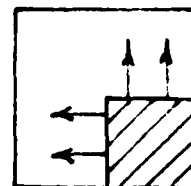




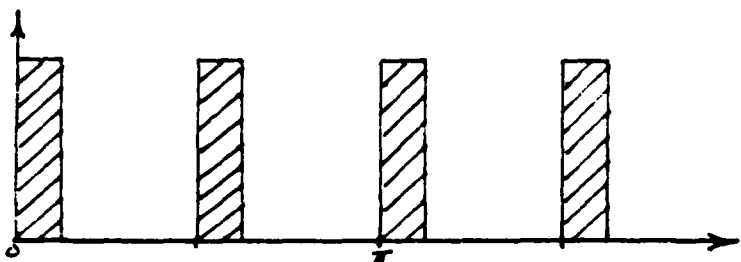
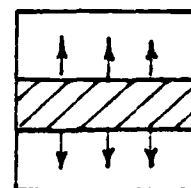
a. Straight Edge,  
Zero Entropy



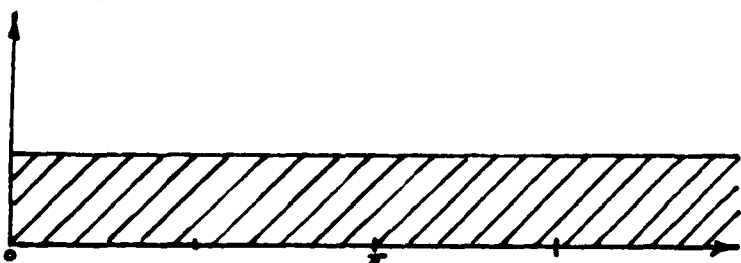
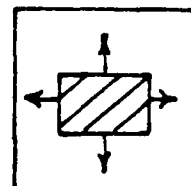
b. Corner,  
Orthogonal Set  
of Gradients



c. Anti-Parallel  
Set



d. Rectangle Pair  
of Orthogonal  
Anti Parallel  
Sets



e. Uniform  
Distribution,  
Maximum Entropy

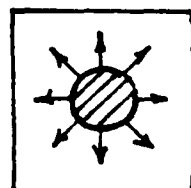


Figure 6. Gradient Histograms

The term  $|\sin \Delta\theta|$  is a relative weighting factor which is maximum for  $\theta_k$  orthogonal to  $\theta_I$  and which diminishes to a zero value as the  $\theta_k$  align with the direction of  $\theta_I$ . To convert these directional bin averages into a single parameter, the following average of  $\omega_I$  was defined:

$$\bar{T} = \sum_K P_K \omega_K = \sum_K \sum_I P_K P_I |\sin(\theta_I - \theta_K)| \quad (43)$$

which is a measure of the orthogonal aspect of the gradients in a sampled window. It is easy to see that  $\bar{T}$  tends to zero for cases like histograms a and c. Windows with larger  $\bar{T}$  values have more orthogonal or omnidirectional attributes which are desirable for selecting localizable and distinguishable pixels. This line of investigation was abandoned because the expressions are computationally cumbersome; however, the concepts discussed above were modified and incorporated into the analytical formalism presented next.

#### 4. ABC'S OF DIRECTIONAL WEIGHTING FUNCTIONS

Several directional weighting functions, such as Equations 42 and 43, evolved from early investigation of histogram statistics. Formally, the definitions of the orthogonal and parallel weighting functions are:

$$w_{\perp}(\theta) = \sum_{i=1}^N \rho_i^2 \sin^2(\theta_i - \theta) \quad (44a)$$

$$w_{\parallel}(\theta) = \sum_{i=1}^N \rho_i^2 \cos^2(\theta_i - \theta) \quad (45a)$$

where  $N$  = Number of gradient samples.

$\rho_i, \theta_i$  = Magnitude and direction of  $i$ th vector (see Equation 38).

Note that these definitions do not depend on the establishment of angular bins such as were used in histogram formalism. It is an easy exercise to derive the following expressions for the weighting functions:

$$2w_{\perp}(\theta) = A - B \sin 2\theta - C \cos 2\theta \quad (44b)$$

$$2w_{\parallel}(\theta) = A + B \sin 2\theta + C \cos 2\theta \quad (45b)$$

where

$$A = \sum_{i=1}^N \rho_i^2 = \sum_{i=1}^N (I_{\alpha}^2 + I_{\beta}^2)$$

$$B = \sum_{i=1}^N \rho_i^2 \sin 2\theta_i = \sum_{i=1}^N (I_{\alpha}^2 - I_{\beta}^2) \quad (\text{same as Equation 33})$$

$$C = \sum_{i=1}^N \rho_i^2 \cos 2\theta_i = \sum_{i=1}^N (I_x^2 - I_y^2)$$

Utilizing  $\sin^2 \Delta\theta$  as a relative weighting factor in  $w_{\perp}$  instead of  $|\sin \Delta\theta|$  as used in Equation 42 provides an analytic function of  $\theta$  with the same A,B,C statistics that were utilized to simplify the derivation of interest measures DIRVAR and EVAR (Equation 33-36).

The weighting function  $w_{\perp}(\theta)$  is a measure of the extent to which the sample gradients are orthogonal to a given direction ( $w_{\perp}(\theta) = w_{\perp}(\theta + \pi)$ ). The tendency of vectors to align themselves in a given direction  $\theta$  is given by the weighting function  $w_{\parallel}(\theta)$ . These dual weighting functions are both obtained for the price of computing the A,B,C statistics. The sum of  $w_{\perp}$  and  $w_{\parallel}$  equals A, which is independent of the relative direction of the gradient vectors. For later reference, we rewrite Equations 44 and 45 in a form that readily depicts the max and min values for these dual functions:

$$2 w_{\perp}(\theta) = A - R \cos 2(\theta - \tau) \quad (44c)$$

$$2 w_{\parallel}(\theta) = A + R \cos 2(\theta - \tau) \quad (45c)$$

where

$$2\tau = \tan^{-1} B/C \quad (46a)$$

$$R^2 = B^2 + C^2 \quad (46b)$$

## 5. ORIENTATION ASPECTS

We now define a new parameter called the orthogonal aspect:

$$T = \sum_i^N \rho_i^2 w_{\perp}(\theta_i) = \sum_i^N \sum_K^N \rho_i^2 \rho_K^2 \sin^2 (\theta_i - \theta_K) \quad (47a)$$

Other expressions are:

$$2T = A^2 - B^2 - C^2 = A^2 - R^2 \quad (47b)$$

$$T = 2 \sum_i \sum_{i < K} (\vec{\nabla} I_i \times \vec{\nabla} I_K) \cdot (\vec{\nabla} I_i \times \vec{\nabla} I_K) \quad (47c)$$

where

$$\vec{\nabla} I_i = [\rho_i \cos \theta_i, \rho_i \sin \theta_i].$$

We see that  $T$  which is a simple function of the gradient statistics  $A, B, C$ , is the continuous analogue to histogram average  $\bar{T}$  (compare Equations 43 and 47a). Equation 47c expresses  $T$  as the sum of the square magnitudes of the cross products between all gradient vector pairs. Recall that the vector cross product between two vectors is another vector,

$$\vec{V}_i \times \vec{V}_K = \vec{W} \quad (48)$$

such that

$$|\vec{W}| = |\vec{V}_i| |\vec{V}_K| |\sin (\theta_i - \theta_K)|.$$

Recall that cross products between parallel or antiparallel vectors vanish. Hence terms with  $i = j$  vanish in Equations 47a and c.

There are two possible definitions for the parallel aspect, both of which are dual to the definitions of the orthogonal aspect given in Equation 47. The first version,  $L$ , is a function of  $A$ ,  $B$ , and  $C$ :

$$L = \sum_i^N \rho_i^2 w_{ii}(\theta_i) = \sum_i^N \sum_K^N \rho_i^2 \cos^2(\theta_i - \theta_K) \quad (49a)$$

$$2L = A^2 + B^2 + C^2 = A^2 + R^2 \quad (49b)$$

$$L + T = A^2 = \left[ \sum_i \rho_i^2 \right]^2 \quad (50)$$

$$L - T = R^2 = \sum_i \sum_K \rho_i^2 \rho_K^2 \cos 2(\theta_i - \theta_K) \quad (51)$$

The second version,  $\Lambda$ , is a counterpart to Equation 47c:

$$\Lambda = 2 \sum_{i < K} (\vec{\nabla} I_i \cdot \vec{\nabla} I_K)^2 \quad (52a)$$

$$\Lambda = L - D \quad (52b)$$

where

$$D = \sum_i \rho_i^4 = \sum_i (\vec{\nabla} I_i \cdot \vec{\nabla} I_i)^2 \quad (53a)$$

is a new gradient statistic. The ambiguity in the parallel aspect definition arises from the fact that unlike a cross product, the dot product of a vector with itself is not identically zero. Equation 51 includes the self products, whereas Equation 52 excludes them. An alternative expression for the sum of self products,

$$D = \sum_i \rho_i^4 = \sum_i (I_a^2 + I_b^2)^2 \quad (53b)$$

illustrates that the computation of  $D$  involves terms  $I_a^2$  and  $I_b^2$  which are used to compute  $A$  and  $B$  so that the extra effort to compute  $D$  is minimal.

Other measures are derivable from the statistics defined above:

$$S = \Lambda + T = A^2 - D = 2 \sum_{i < K} \rho_i^2 \rho_K^2 \quad (54)$$

$$Q = \Lambda - T = R^2 - D = 2 \sum_{i < K} \rho_i^2 \rho_K^2 \cos 2(\theta_i - \theta_K) \quad (55)$$

where

$$-1 \leq Q/S \leq 1.$$

The following three ratios may be interpreted as averages of various trigonometric functions of  $\Delta\theta_{iK} = \theta_i - \theta_K$  with the product of magnitudes  $(\rho_i \rho_K)^2$  utilized as the weighting parameter:

$$\begin{aligned} Q/S &= \langle \cos 2\Delta\theta \rangle = \cos 2\bar{\Delta\theta} \\ \Lambda/S &= \langle \cos^2 \Delta\theta \rangle = \cos^2 \bar{\Delta\theta} \\ T/S &= \langle \sin^2 \Delta\theta \rangle = \sin^2 \bar{\Delta\theta} \end{aligned} \quad (56)$$

where

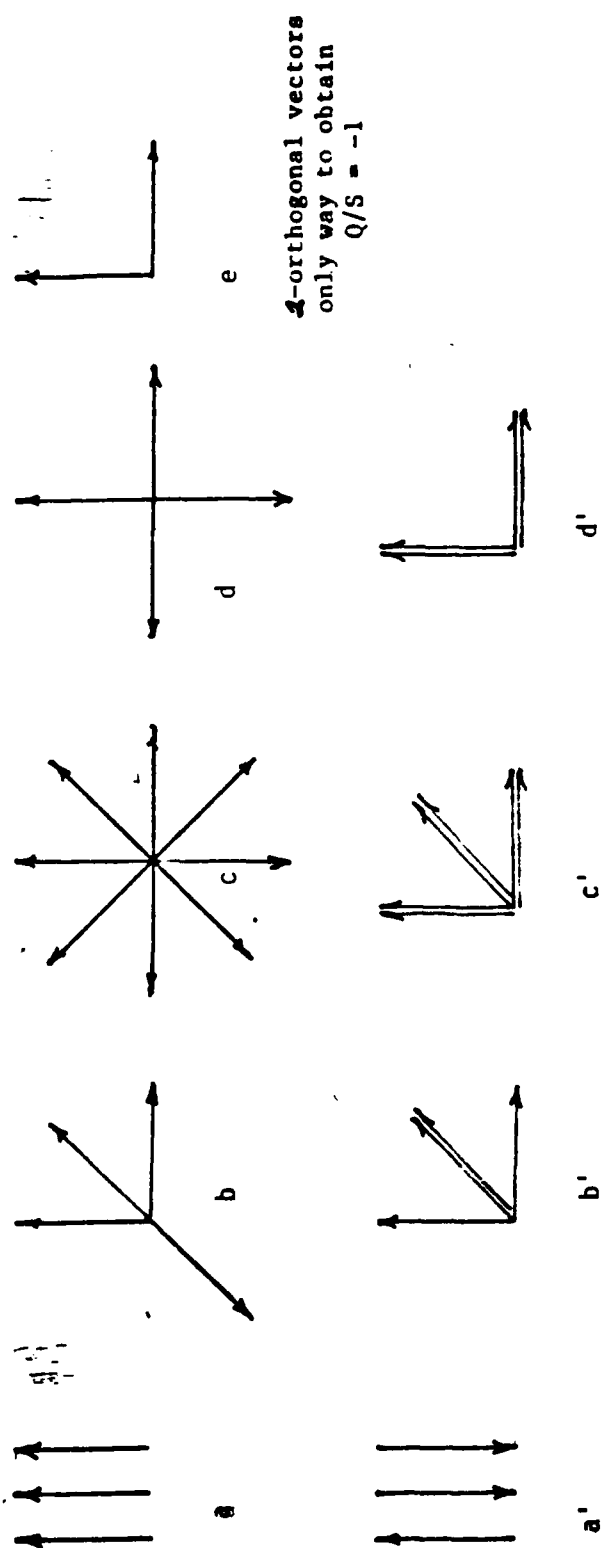
$$0 \leq \bar{\Delta\theta} < \pi.$$

Figure 7 depicts various sample sets of unit vectors and corresponding statistical measures, which include limits for  $Q/S = \pm 1$ . Statistics derivable from A, B, C, and D, which are sums of terms with even powers, do not distinguish between parallel and antiparallel vectors as depicted by seemingly different sets with identical even order statistics.

The mean gradient vector ( $G, \bar{\phi}$ )

$$G \cos \bar{\phi} = \langle I_a \rangle = \frac{1}{N} \sum_i \rho_i \cos \phi_i \quad (57)$$

$$G \sin \bar{\phi} = \langle I_b \rangle = \frac{1}{N} \sum_i \rho_i \sin \phi_i \quad (58)$$



	L	A	T	S	Q	Q/S	$\lambda\theta = \lambda \cos^{-1}(Q/S)$
a, a'	9	6	0	6	6	1	0°
b, b'	10	6	6	12	0	0	45°
c, c'	32	24	32	56	-8	-1/7	50°
d, d'	8	4	8	12	-4	-1/3	54°
e	2	0	2	2	-2	-1	90°

Figure 7. Gradient Statistics: Examples use unit vectors in eight principle directions

is useful for characterizing the parallel/antiparallel aspects of the gradient sets. The direction,  $\bar{\phi}$ , of the mean gradient provides an invariant reference direction. The projection of the mean gradient vector onto any direction,  $\theta$ , is given by

$$M(\theta) = \sum_i \rho_i \cos(\theta_i - \theta) / N = G \cos(\bar{\phi} - \theta) \quad (59)$$

$$G^2 = \frac{N}{\sum_K \rho_K} M(\theta_K) / N = \frac{\sum_K \sum_i \rho_i \rho_K \cos(\theta_i - \theta_K)}{N^2} \quad (60)$$

Equations 59 and 60 demonstrate more explicitly the dependence of  $G$  on the extent to which the vector pairs are parallel or antiparallel. Note that while the mean gradient vector is invariant, the vector components depend on the coordinate system used to define them. If  $\langle I_x \rangle$  and  $\langle I_y \rangle$  were used in Definition 57,  $G$  would be the same and the reference angle,  $\bar{\phi}$ , would be changed by  $\pi/4$  in the rotated coordinates.

## 6. NEW INTEREST MEASURES

In this segment, we shall explore interest measures utilizing the gradient formalism already introduced. The modified directional statistics  $D(i)$  utilized in the definitions of DIRVAR and EVAR (see Equations 34-36) can be expressed in terms of (see Equation 44b):

$$\begin{aligned} 2D(1) &= w_l(3\pi/4) = A + B \\ 2D(2) &= w_l(3\pi/4) = A - B \\ 2D(3) &= w_l(\pi/2) = A + C \\ 2D(4) &= w_l(0) = A - C \end{aligned} \quad (34a)$$

so that

$$2 \text{ DIRVAR} = \text{MIN} [w_l(0), w_l(\pi/4), w_l(\pi/2), w_l(3\pi/4)] \quad (35a)$$

$$2 \text{ EVAR} = \text{VAR} * \text{MIN} \left[ \frac{w_l(0)}{w_l(\pi/2)}, \frac{w_l(\pi/2)}{w_l(0)}, \frac{w_l(\pi/4)}{w_l(3\pi/4)}, \frac{w_l(3\pi/4)}{w_l(\pi/4)} \right] \quad (36a)$$



Note that these measures depend on  $w_l(\theta)$  evaluated in four principal directions, which in turn depend on the particular coordinate system in use. The continuous and rotationally invariant analogues to these equations are:

$$2 \text{ DIRVAR}' = \text{MIN } [w_l(\theta)] = w_l(\text{min}) = A - B \quad (61)$$

$$2 \text{ EVAR}' = \text{VAR} * w_l(\text{min})/w_l(\text{max}) = \text{VAR} (A - R)/(A + R) \quad (62)$$

where the angle between  $w_l(\text{max})$  and  $w_l(\text{min})$  is  $\pi/2$ . Here we have replaced MX in Equation 38 and 39 with  $R = \sqrt{B^2 + C^2}$ .

Reexpressing the orthogonal aspect in terms of  $w_l(\text{max})$  and  $w_l(\text{min})$  yields

$$T = A^2 - R^2 = w_l(\text{max}) * w_l(\text{min}) \quad (63)$$

which suggests that  $T$  be considered as an interest measure for point selection. The interest measures defined in the last three equations are all analytic, rotationally invariant, and proportional to  $w_l(\text{min})$ . The relative merits of the various new measures introduced in this paper await future computer simulation experiments using different types of imagery.

## 7. SECTION SUMMARY

In classical pattern recognition, one extracts a set of features (feature vectors) which is used to distinguish between predetermined classes of subimages or textured patterns. Effective features provide good separation of the predetermined classes in the multidimensional feature space. In image analysis (or image understanding), one extracts locally derived features (image primitives) and attempts to integrate the simpler primitives into more complex and globally extended structures, which are meaningful to the applications at hand. The basic primitives are edges, texture, contrast (all of which are to some extent characterized by gradients), and color.

Sets of image primitives or statistical features may be relatively easy to compute as locally derived functions of the A, B, C, and D statistics, along with first order statistics ( $G, \phi$ ). Dual directional statistics have been discussed which measure relative orientations between pairs of gradient vectors. A complete evaluation of

these aspects and other rotationally invariant functions will require further experimentation and computer analysis of digitized images. The following brief summary of application areas for gradient statistics includes new areas for further research:

1. Point selection (interest) operators. These were fully discussed, including old and new definitions.
2. Local edge detectors. By definition these detectors are functions of local gradients. See the discussion on the Sobel edge detector in this section.
3. Image segmentation. To formulate more extensive and unified image features, images are segmented into regions with uniform or similar features. Color is a good feature to use. The use of uniform gradient statistics may also apply to the segmentation problem.
4. Feature detection. A preliminary survey of the values of T and L for small binary images indicates that T is sensitive to corners and crosses, while L is sensitive to straight lines and fibrils (antiparallel sets of lines). This is encouraging; however, to fully explore the discriminating power of the directional aspects, interactive experiments with grey level images are required. These operators can be used to cue areas of potential interest.
5. Textural analysis. Although not discussed in this paper, this subject has been an active research topic for the last 25 years or so; hence there is an extensive literature available on it. The techniques for analyzing and synthesizing texture in images are generally divided into structural and statistical methods. The first approach looks for repetitive patterns of image primitives such as lines, simple shapes, etc. The statistical approach utilizes statistical parameters which characterize the distribution patterns and spatial relationships between pixel gray levels. Some techniques use extensive histograms called Spatial-Grey-Level Dependence Matrices (Reference 14, p. 186) which quantify differences not just of adjacent pixels, but of pixels with different relative separations. Other techniques analyze polar histograms obtained from partitioning the Fourier power spectrum. These techniques require extensive computations. The relatively simple gradient features outlined above may provide simpler alternative approaches for working with texture.

## SECTION V

### ADAPTIVE IMAGE MATCHING ALGORITHM

The problem of computing complete image-to-image mapping in real time is unsolved. In fact, computing a complete disparity field is not well understood. Point matching algorithms are robust and computationally extensive, but cannot solve the complete mapping problem by themselves. One attempt to achieve complete mapping uses spatial-temporal gradient analysis with a smoothness constraint (see Section II). The smoothness constraint results in iterative equations to "fill in" uniform intensity regions. The problem with this approach is that the fill-in process needs to start from high contrast boundary regions in which the smoothness constraint is weakest. The experimental approach as presented here also addresses the complete mapping problem.

An adaptive binary image matching algorithm may be characterized as a blob matcher. In this approach, a uniform region is not viewed as something to be filled in or as a region devoid of matchable points, but as a unique feature that can be isolated in separate images and compared. In viewing aerial photographs, a lake will stand out as a distinctive uniform area, or blob, which is a key feature in one's initial registration of multiple images. Unfortunately, there are not enough of these naturally occurring blobs to provide a complete disparity map. The proposed approach investigates means for generating artificial blobs. In this way, the blob matching process can proceed until a completely integrated picture of the disparity field is accumulated.

The problem of generating artificial blobs is equivalent to transforming grey-level images into binary images. In binary images, the concept of a blob is obvious; i.e., a blob is an island of black pixels in a sea of white pixels. There are an infinity of such image transformations. Generally, binary images are achieved with the following two steps: transforming the grey-level image into a different image representation, and thresholding the new image representation. All conceivable transformations are not equally suitable. Sought are binary mapping transformations which have high expectations that corresponding physical elements will map into corresponding blobs. For example, selecting the brightest regions (or darkest) in an image will generally be expected to correspond to the same surface elements in another image.

An adaptive image matching system, together with experimental support facilities, is depicted in Figure 8. Inputs to the system are two images,  $I$  and  $I'$ , and the output is a complete disparity vector field,  $\vec{D}$ , which maps the corresponding points from one image to the other. The input images are subjected to a sequence of image transformations,  $T(I)$ , each of which may be transformed into one or more binary images by thresholding. Each pair of binary images  $B(I)$  and  $B(I')$  is compared. The relative displacement of elements in the images provides information about the disparity field,  $\vec{\Delta D}$ , over the area of the blobs, thus providing a piece of the disparity puzzle. Some of the pieces, however, may overlap and could have somewhat different estimates for the disparity field in these overlapping regions. All partial results are accumulated and integrated in the final block in Figure 8 which also provides executive control over the sequencing of transforms, thresholding, and interfacing with the interactive workstation containing an extensive software library.

The functional blocks in Figure 8 should be viewed as software modules for near term simulation and hardware modules for a future prototype real-time image processor. A realistic research program should constantly attempt to match simulated algorithms with identifiable hardware technology. The scope of this experimental approach is very broad, so that a hardware implementation of successful simulated algorithms could evolve into an experimental test bed for promising image processing technology.

A workstation, built around a computer such as the VAX-11/780, provides a general purpose research and development tool for various laboratory projects, as well as the basis for the matching system. Special care in designing efficient data structures and management procedures is needed because of the large flow of data involved in accumulating the complete disparity field. The library should be stocked with published algorithms and available software packages. This includes a package to simulate cellular logic or systolic arrays to explore morphological image analysis for comparing binary images.

The initial set of transforms used in the matching experiments will consist of commonly used algorithms utilized in pattern recognition and image understanding. These tried and tested algorithms can be expected to highlight physically meaningful attributes, such as texture measures and gradient statistics. Color, of course,

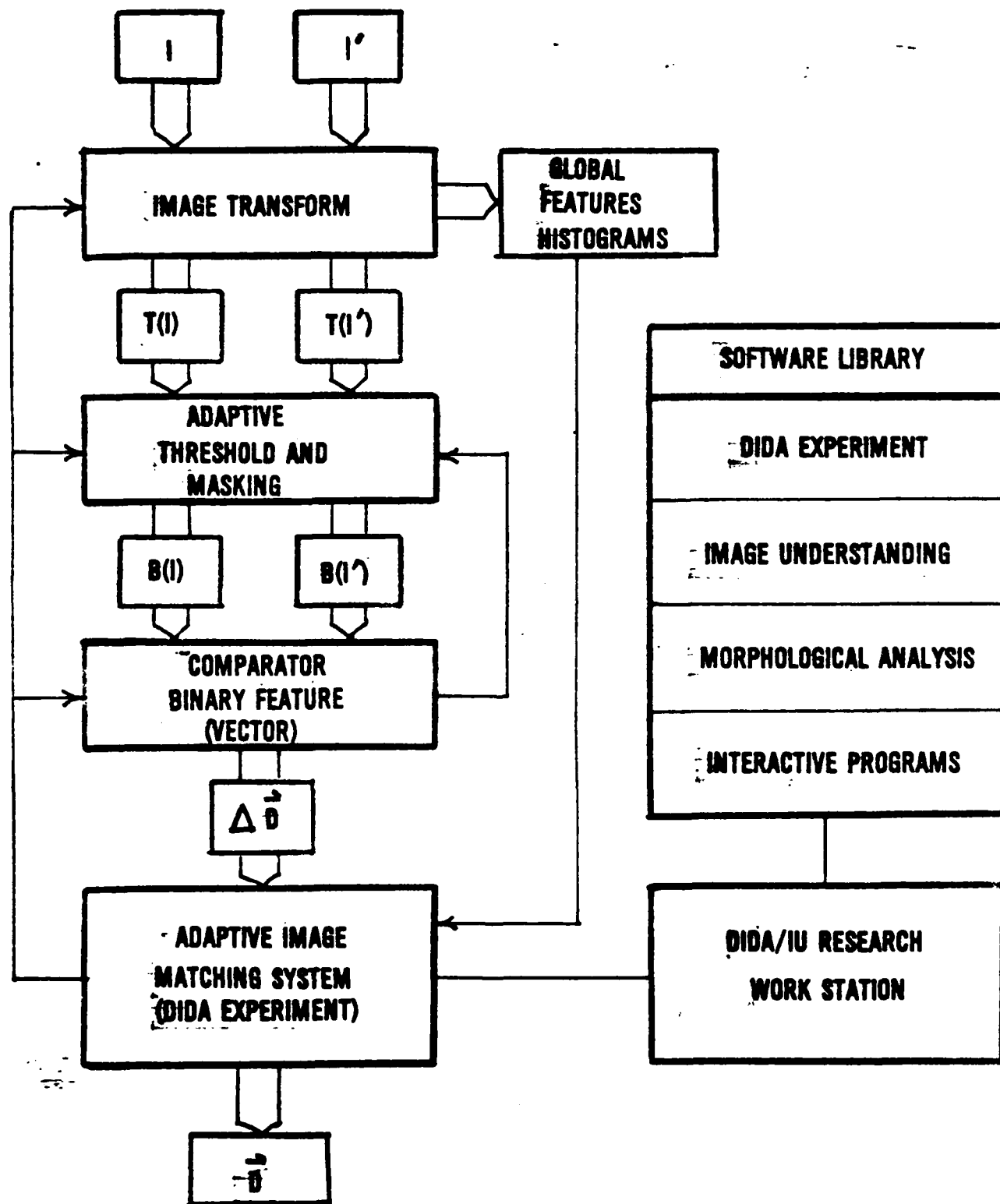


Figure 8. Adaptive Image Matching System

would be an excellent feature to use if available. By keeping histories of performance, each transform can be evaluated and indexed. Given the ability to measure performance, automatic means for learning new and useful transforms can be developed for enhancing the system's repertoire. More sophisticated learning algorithms can learn to select a sequence of transforms to match a priori or contextual knowledge about the images being matched. Most of the image transforms will be local neighborhood or window functions to facilitate real-time implementation schemes. A special module is envisioned for extracting global features and histograms from the images, in order to supply the contextual information for the adaptive transform generation.

The feed back path from the comparator to the threshold module would be used to vary the threshold settings in order to achieve convenient blob density and sizes. The masking ability would be used to perform blob matching or to concentrate attention in those regions which are sparsely represented in the disparity field accumulator. It may be feasible to generate a confidence measure for each entry in the accumulator. Similar estimates derived from different transforms should increase these confidence measures. Different estimates will require arbitration.

The use of binary image matching has been presented as a promising alternative for accumulating complete disparity maps. Image matching systems of the future will be hybrid systems incorporating a variety of matching algorithms such as reviewed in Section II. The velocity constraint equation ( $\vec{\nabla} I \cdot \vec{v} + \partial I / \partial t = 0$ ), for example, could help delineate corresponding blobs and provide independent measurements for cross-checking the accumulating matching results. The technical disciplines needed to implement real-time systems include image understanding, artificial intelligence, and dynamic image analysis, plus super computers for image processing. All of these disciplines are heavily supported by DOD. One of the original objectives for creating the DIDA program was to help establish a center for machine perception. This center would in turn provide an alternative and effective means for harvesting the results of DOD's research investment.

# REFERENCES

1. R. Jain and H. Nagel, "On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes," IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-1, 206 - 214, April 1979.
2. W. Martin and J. Aggarwa, "Survey of Dynamic Scene Analysis," Computer Graphics and Image processing, 7, 3, June 1978.
3. R. Jain, "Dynamic Scene Analysis Using Pixel-Based Processor," IEEE Computer, p. 12, August 1981.
4. R. Jain, "Extraction of Motion Information from Peripheral Processes," IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-3, 5, 489 - 503, Sept 1981.
5. E. Horn and B. Schunck, "Determining Optical Flow," Proceedings, DARPA Image Understanding Workshop, 144-156, April 1981.
6. B. Horn and B. Schunck, "Constraints on Optical Flow Computation," Proceedings, IEEE Conference on Pattern Recognition and Image Processing, 205 - 210, August 1981.
7. W. Thompson and J. Kearney, DIDA - Dynamic Image Disparity Analysis, Avionics Laboratory, WPAFB, Ohio, AFWAL-TR-83-1035, December 31, 1982.
8. S. Barnard and W. Thompson, "Disparity Analysis of Images," Trans. Pattern Analysis and Machine Intelligence, PAMI-2, 4, 333 - 340, July 1980.
9. W. Thompson and S. Barnard, "Lower-Level Estimation and Interpretation of Visual Motion," IEEE Computer, August 1981.
10. H. Moravec, "Visual Mapping by a Robot Rover," Proc. 6th Int. Joint Conf. on Artificial Intelligence, 598 - 600, August 1979.
11. H. Moravec, "Towards Automatic Visual Obstacle Avoidance," Proc. 5th Int. Joint Conf. on Artificial Intelligence, 584, August 1977.
12. M. Hannah, "Bootstrap Stereo" Proceedings, DARPA Image Understanding Workshop, 201 - 208, April 1980.
13. M. Hannah, Passive Image Navigation, Final Technical Report, DARPA Contract F33615-78-C-1612, Lockheed Palo Alto Research Laboratory, LMSC D67313, Section 3, December 1980.
14. D. Ballard and C. Brown, Computer Vision, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982, pp. 166 - 192.

DATE  
FILMED  
0-8